FEE

Big Data for Operator Support in Chemical Plants

Introduction

Deutsches Zentrum
für Luft- und Raumfahrt e.V.
Projektträger im DLR

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Chemial Industry – A case for Big data

Big Data

- **High Volume, e.g.:**
  - > 300 GB measurement data p.a. in a single refinery
  - 400 GB alarms & events p.a. in a single petro-chemical plant
- **High Velocity, e.g.:**
  - 66.000 sensor with sampling rates between 1s – 60s
- **High Variety, e.g:**
  - Time-Series, Log files, unstructured text, video data
- **Low Veracity, e.g:**
  - time-synchronisation, faulty measurement, missing data

Chemical Industry – A challenge for Big Data

- Challenging problems für data analytics – more machine learning then simple statistics
- Data collection processes not optimized for Big Data Analytics
- High efforts for data exploration due to data silos with unstructured and inconsistent references
- High efforts for data-preperation and cleansing due to interrelations unknown to the data analyst

Project Overview

Heterogeneous Mass data

Big Data Analysis

Support System

Operator

ABB

BASF

rapidminer

INEOS THE WORD FOR CHEMICALS

PCK Raffinerie GmbH

TECHNISCHE UNIVERSITÄT DRESDEN

UNIKASSEL VERSITÄT

Objective: Operator Support functions

- Early Warnings
- Ad-hoc Analysis
- Decision Support

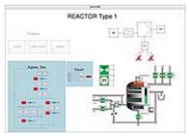
Approach: Integrated Analysis of all plant data

- Measurements, engineering data, electronic shift books,…

Research Topics

- Algorithm development
- Indexing of and search in process data
- Integration into real-time plant operation
- Big data technologies and architecture
- User Centered interaction concepts

FEE – Data and System Landscape



Engineering Data

Laboratory Data

Alarms & Events

Process Measurements

Asset Data

Operation Manuals

Big Data Platform

Assistence Systems
- Early Warnings
- Anomaly Detection & Scenario Search
- Ad-hoc-Analysen & Prognosis

Early Warning

Interactive Usage

Learning from History

Production Plant

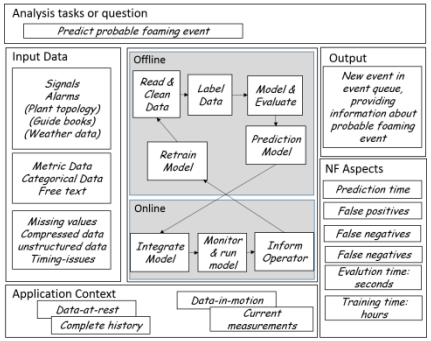
Digital Shift book

operators

FEE – Development Appraoch

1. Scenario Identification

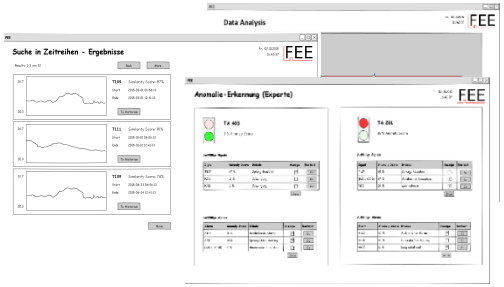
2. Paper Prototypes



3. Analysis Workflows & Non-Functional Requirements

4. PoC Data Analytics

6. Integration and Deployment



5. Refined Mock-ups

FEE

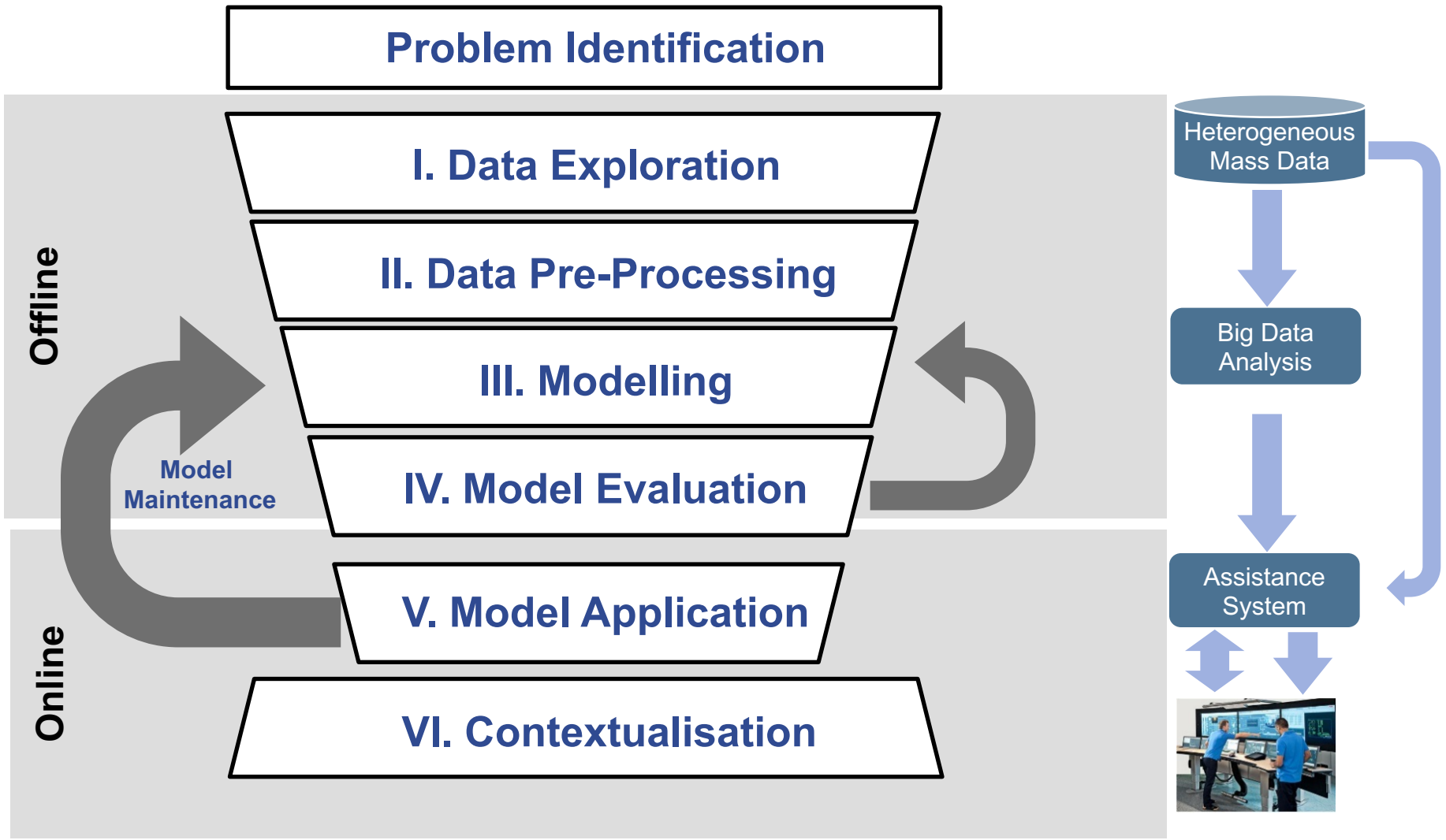
Big Data for Operator Support in Chemical Plants

Szenario – From Big Data to Smart Data

Deutsches Zentrum
für Luft- und Raumfahrt e.V.
Projektträger im DLR

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Life Cylce of Operator Support Function



Scenario: Prediction of Foaming Event

Current State:

Who:

- Operators in control room and in the field

What:

- Foaming in a process column results in increase pressure and risk of spillover. Anti-foaming agent needs to be added manually.

How:

- Monitoring relevant signals in the control room

Problems:

(1) Risk of not recognizing foaming early enough
(2) Foaming is a fast process – actions are always taken under time pressure
(3) Unexperienced operators might not recognize the situation or do not know how to react

Desired State:

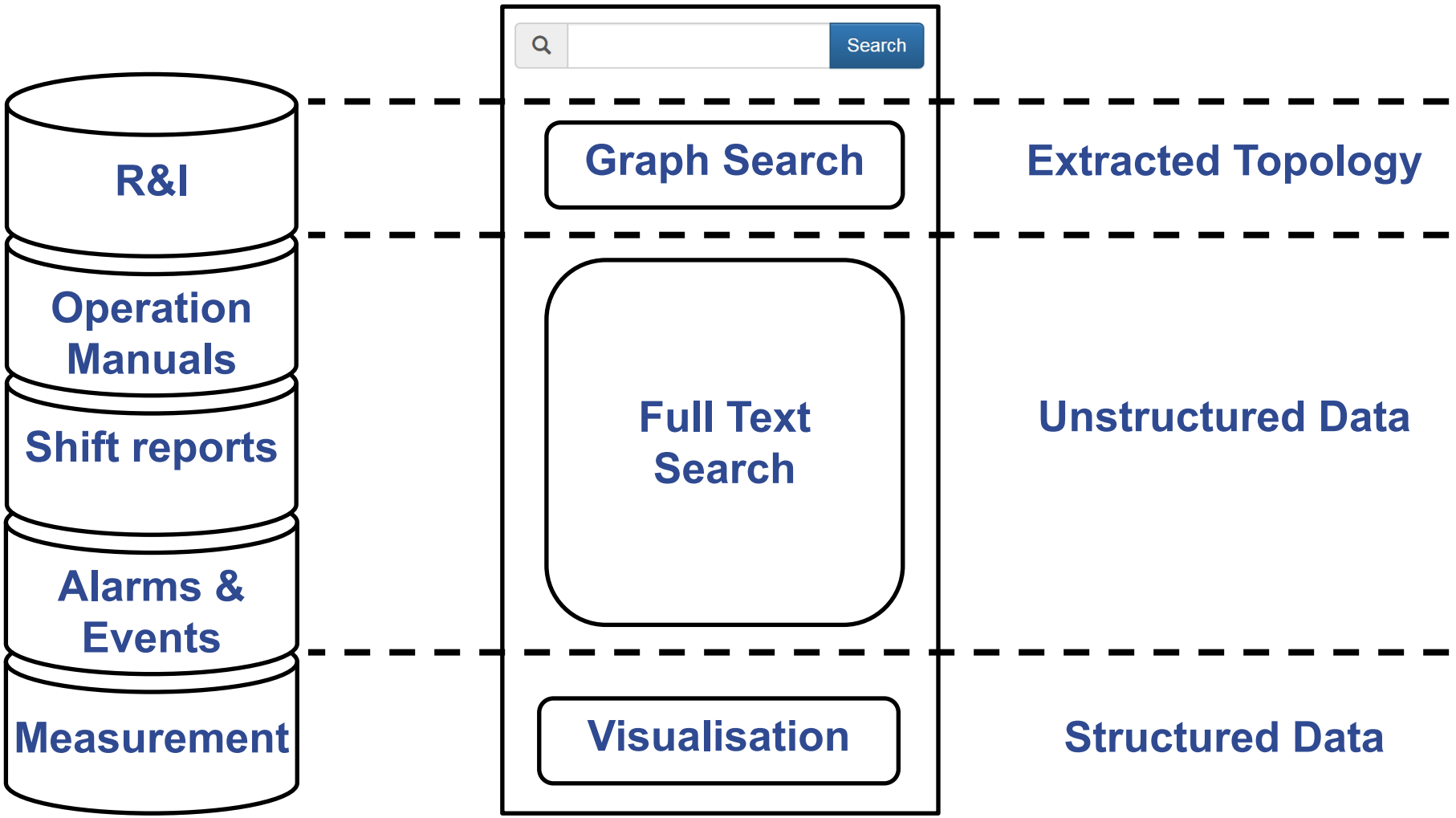
FEE Support:

Early information about certain or probably foaming in the new future.

Desire:

(1) Timely information – latest 30 minutes before the foaming
(2) Clear and specific instructions, no need for diagnostics activities
(3) High prediction rate, few false alarms

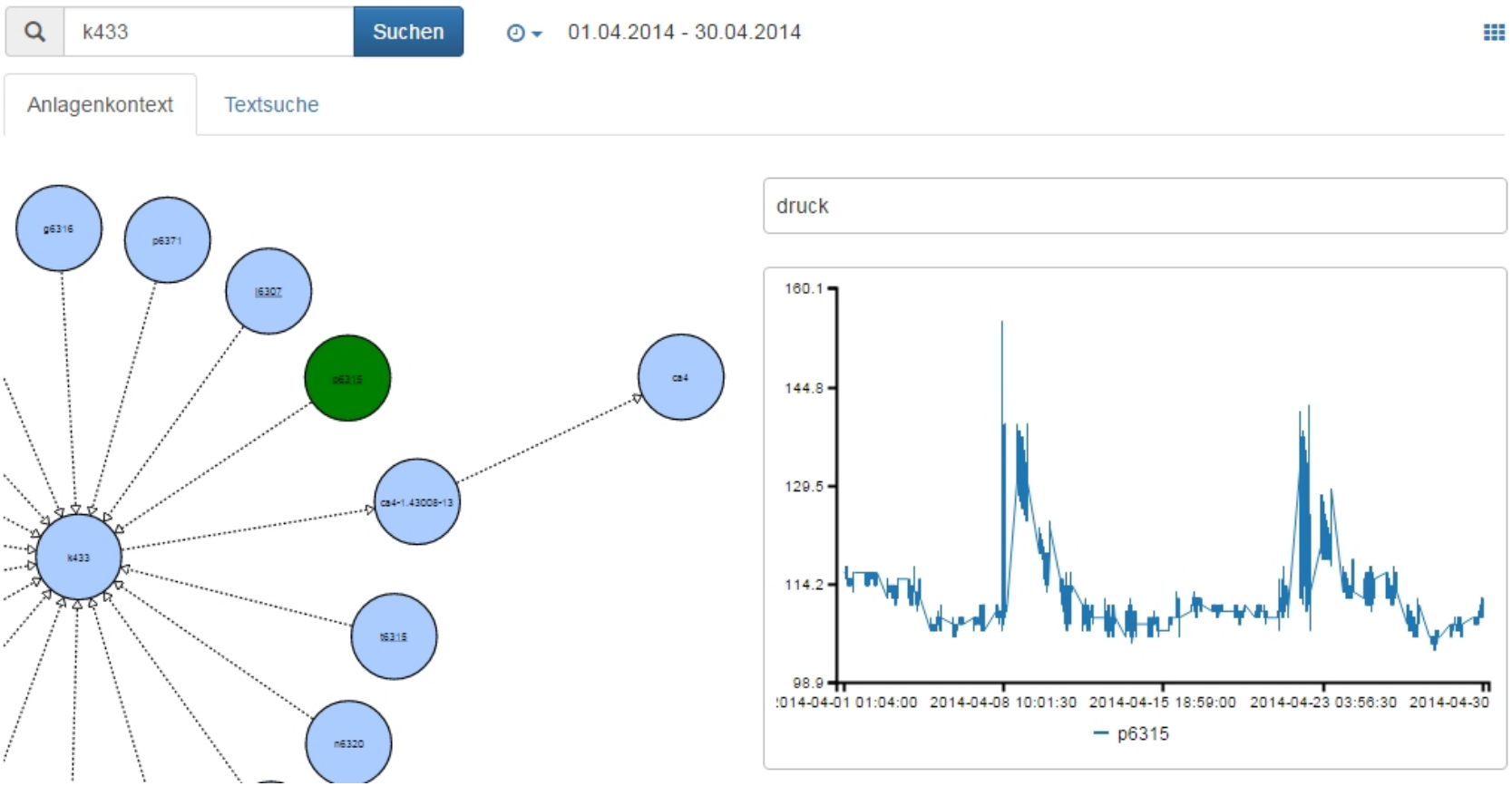
Hybrid Data Exploration



Full Text Search across all Data Sources

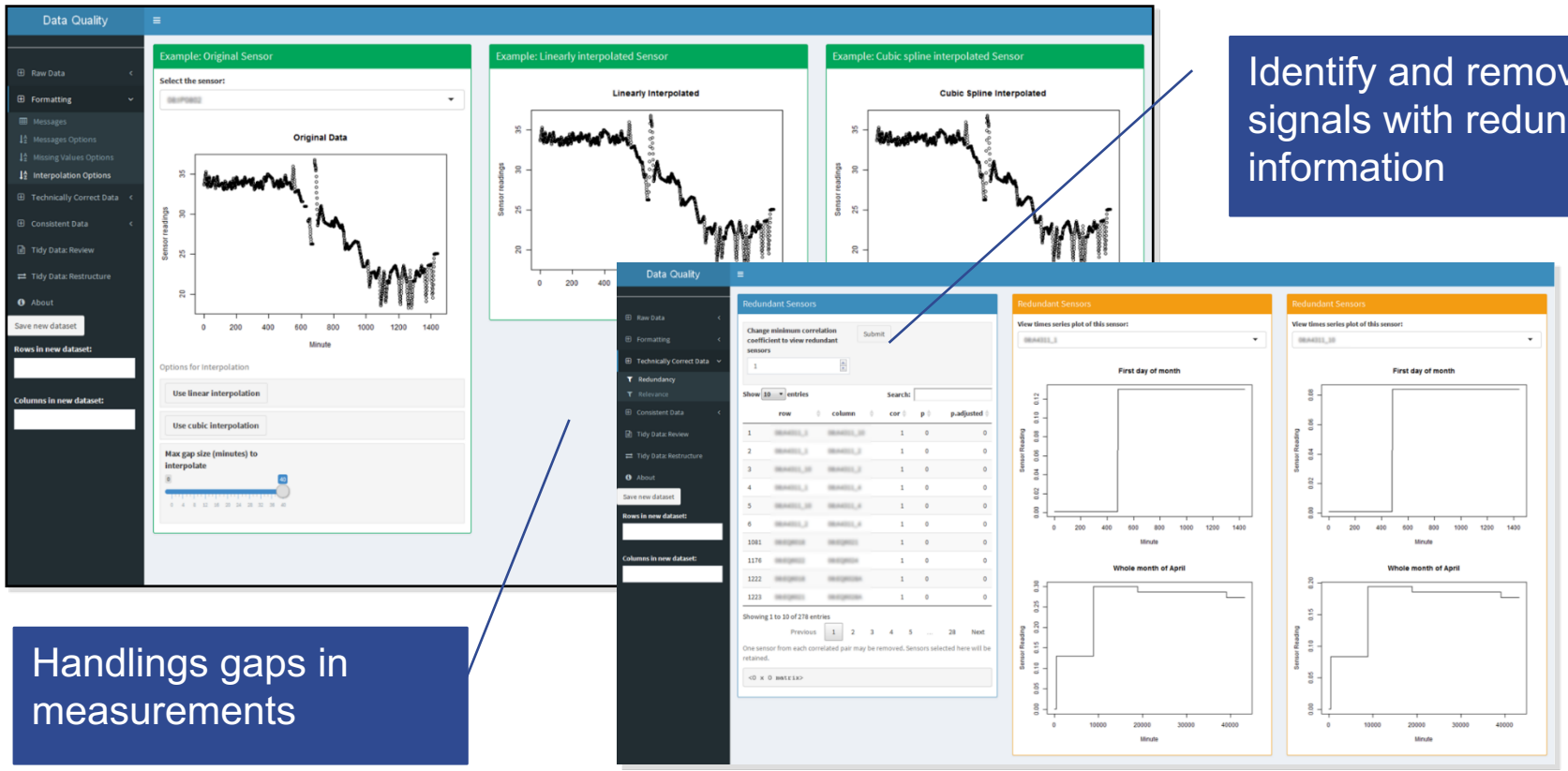
Simple access to data by full text search

Topology Browsing



Graphical Exploration of data based on derived plant topologies

Tool supported Data Exploration

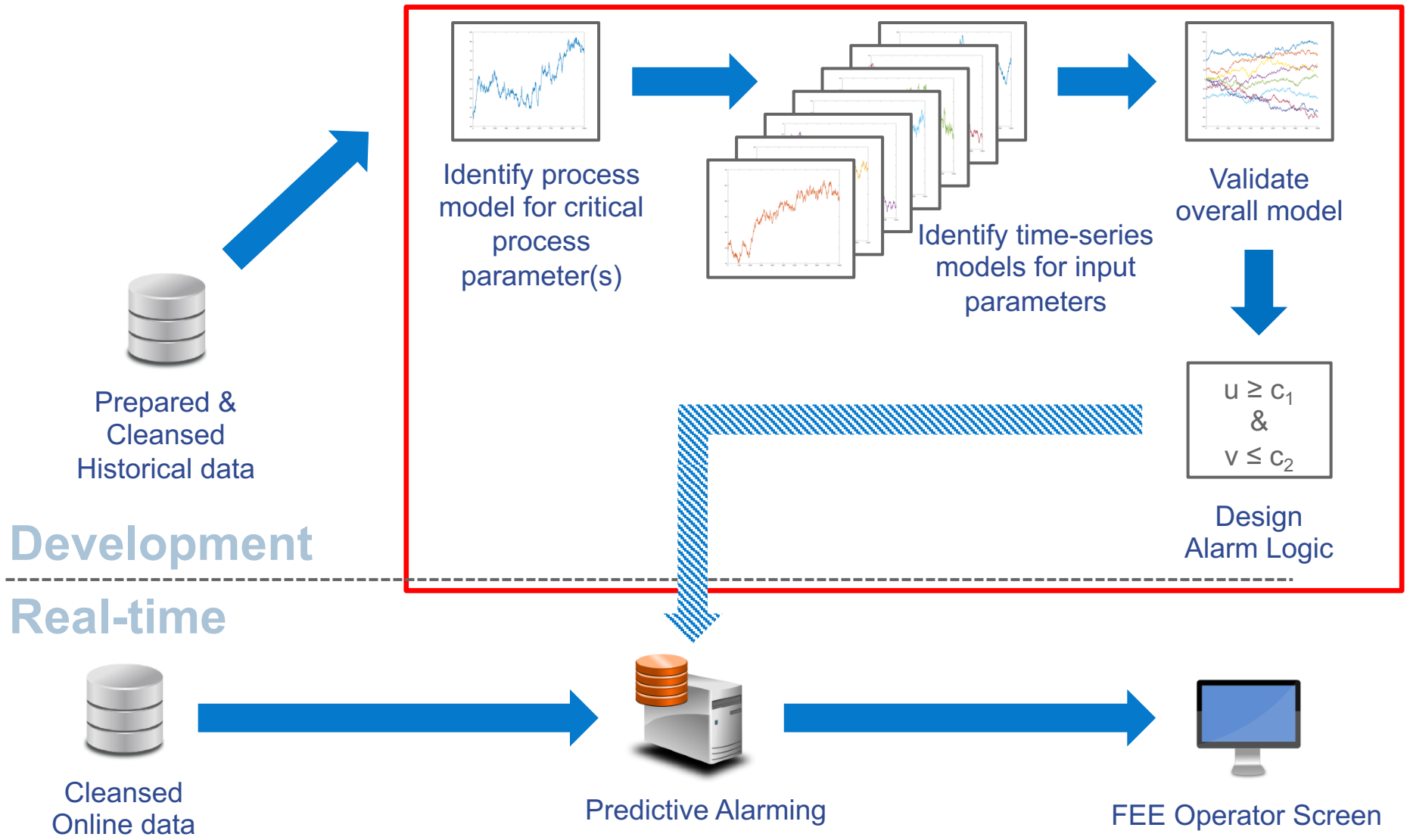


Identify and remove signals with redundant information

Handlings gaps in measurements

Speed-up typical data cleansing & selection tasks

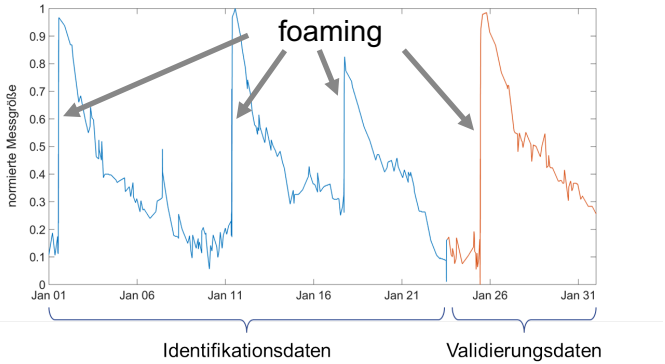
Modelling, -validation and application



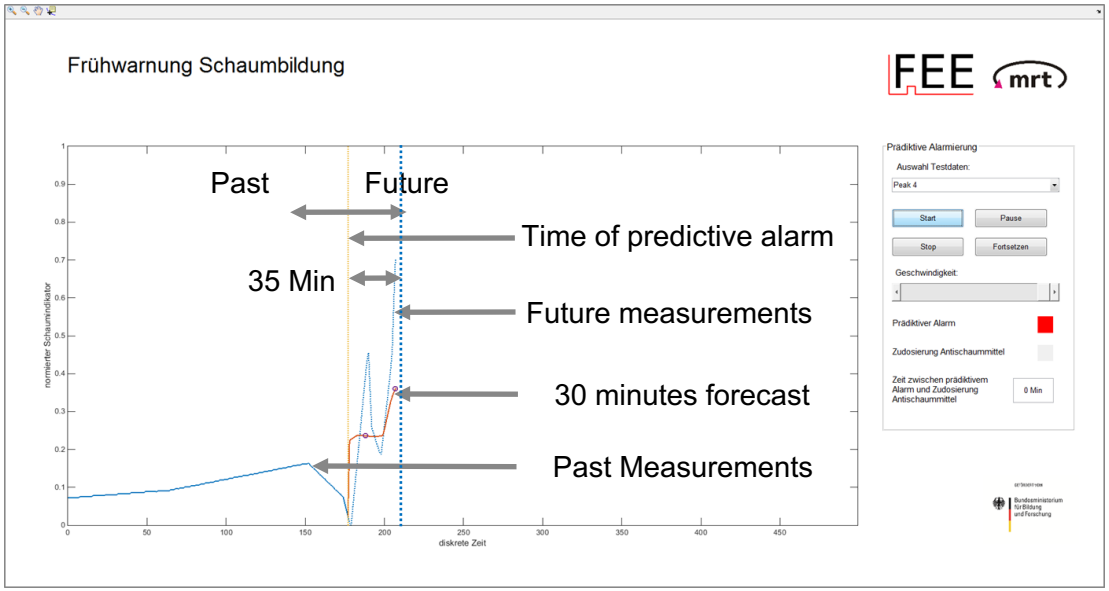
Case study: foaming detection in SCOT plant

- Automated selection of significant input signals and model terms for ARX process model
- Automated selection of significant model terms for AR time-series models
- Overall validation by iterative multi-step prediction
- Simple alarm logic on predicted signals (threshold for signal amplitude and signal gradient)

- Critical signal:



- Predictive Alarming from Engineering Perspective



- Sampling: 1 Min
- Measurements per signals: 44641
- Potential Input Signals: 29
- Significant Input signals : 7
- Timeliness of predictive alarm: 35 minutes

FEE

Big Data for Operator Support in Chemical Plants

Scenario – Anomaly Detection: Big Data for rare events

DLR Deutsches Zentrum für Luft- und Raumfahrt e.V. Projektträger im DLR

GEFÖRDERT VOM

Bundesministerium für Bildung und Forschung

Scenario: Detection of supicious signal paths

Current State:

Who:

- Operator in the control room (and process engineers)

What:

- Monitoring of the plant in ‚calm' situations

How:

- Browsing operator screens and trend display for suspicious signals
- Is only done in ‚calm' situation without stress

Problems:

(1) Risk to simply overlook a suspicious signal
(2) Monitoring without broad coverage in stressful situations
(3) Difficult for unexperienced operators to judge the , suspiciousness' of signals

Desired State:

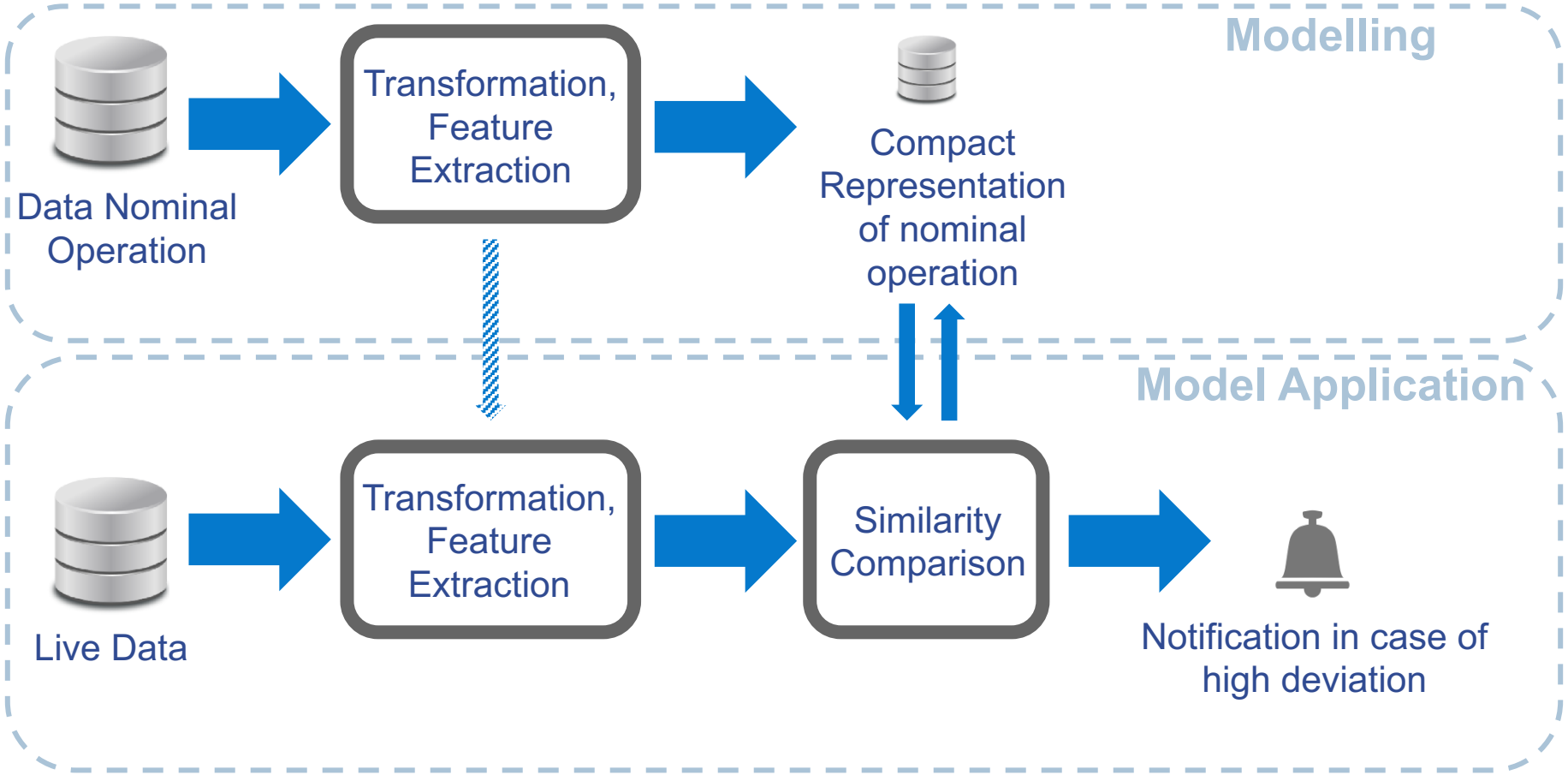
FEE Support:

Identify suspicious signals and providing relevant data for diagnosis

Desire:

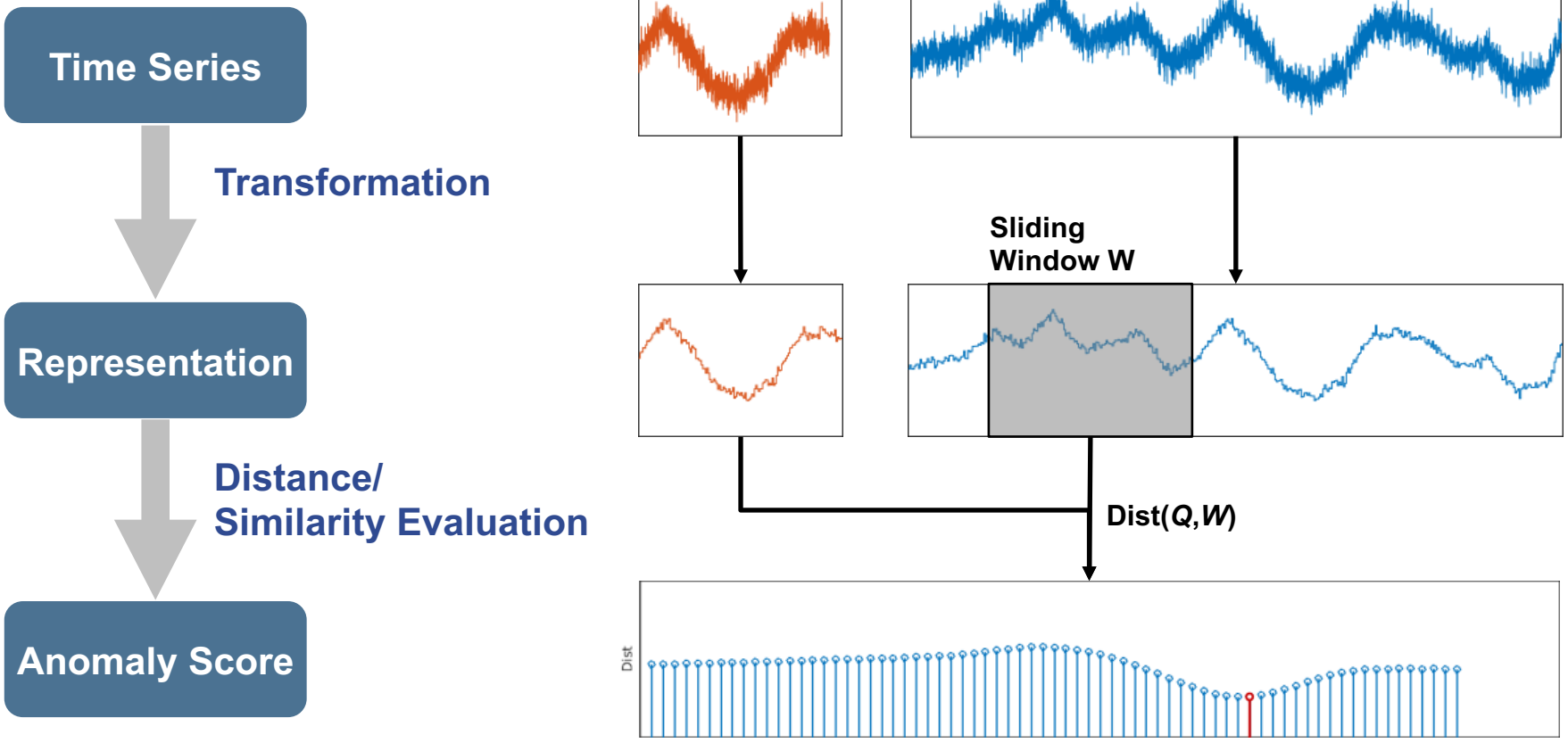
(1) Fast visual impression on abnormalities in the process
(2) Put into context to historical ‚normal' and ‚abnormal' signal paths
(3) Providing extended context (relevant alarms, operator notes, documents)

Anomalie-Erkennung
Analyseworkflow



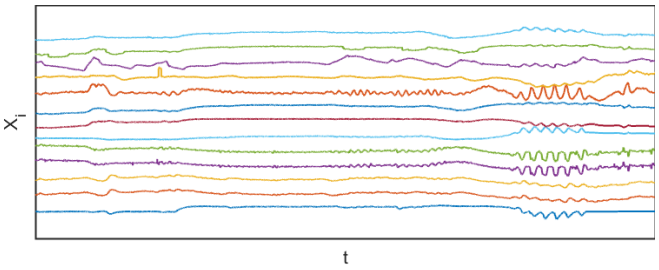
Subsequence Matching basierte Anomaliedetektion

The distance between a live data time-series and the most similar subsequence from historical database is used to calculate the anomaly score.

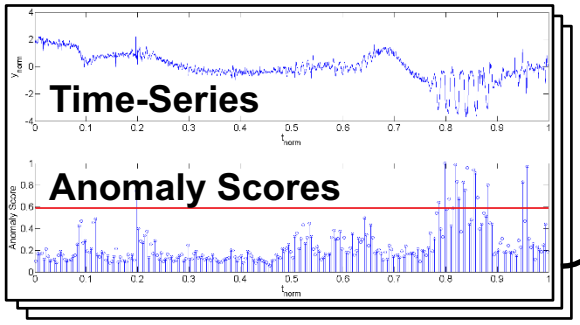


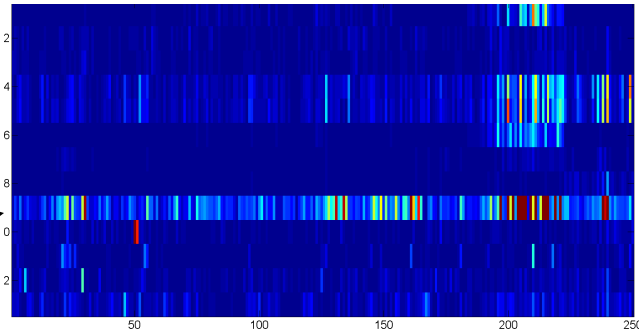
Case Study – Oscillation Detection

- Continuously operated butadiene plant
- One (known) singular anomaly
- **High Data Volume:** ~1000 measuring points with sampling rate of 1 minutes over two years
- **Heterogeneous:** Pressure, flows, levels, analyzer, temperatures, varying compression over time and different from time-series to time-series
- **Nonstationary: Freq**uent load changes
- **Data Selection**:
  - Data selection without expert knowledge: Elimination of redundant and constant time-series to 104 measuring points
  - Data selection by expert knowledge: 13 measuring points (shown)



- Visualization of calculated anomaly scores in a heat map





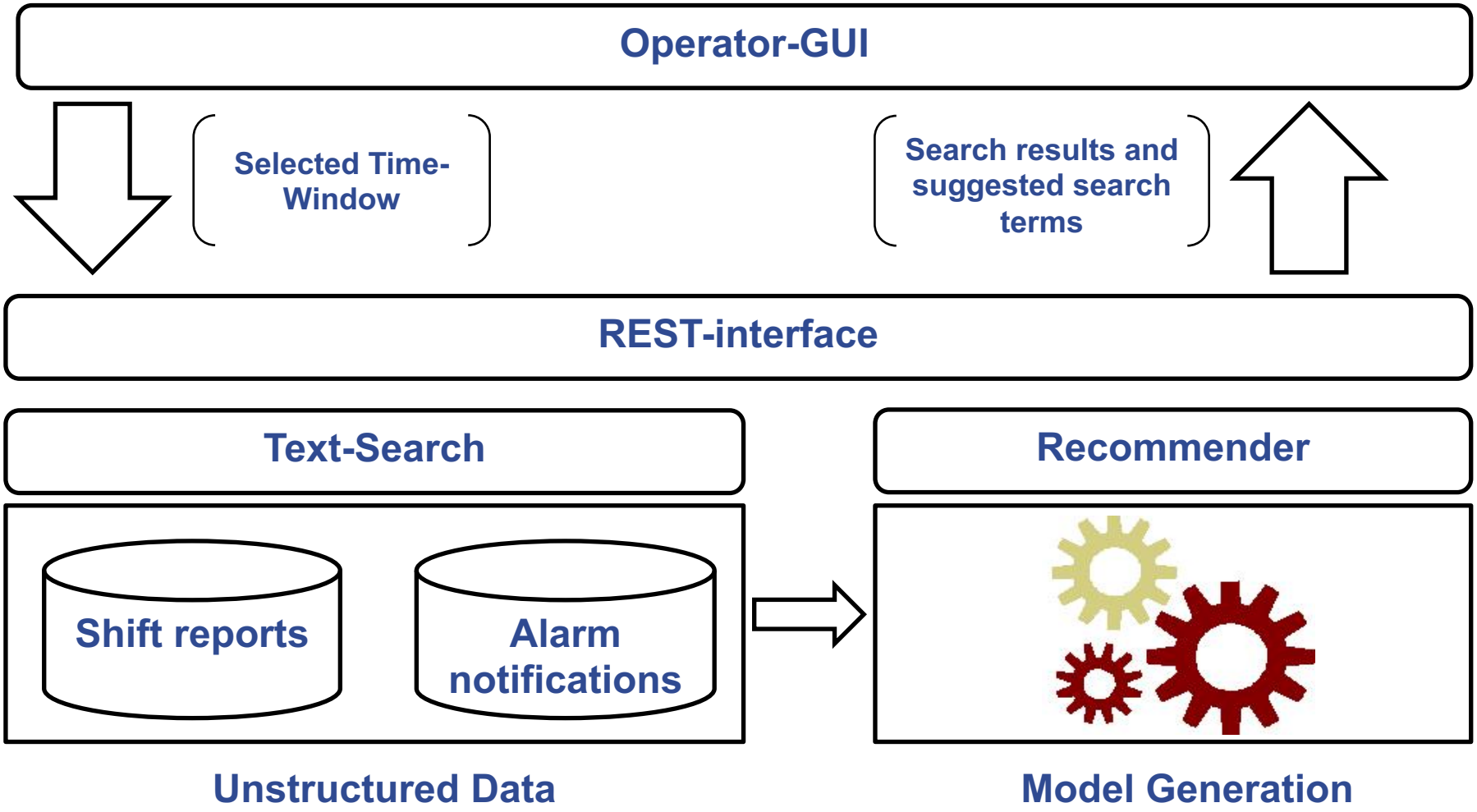
Operator support by Search Term Suggestion

- Information available in unstructured formats

- Objective: Support operator in finding information by suggestion of context-sensitive search terms

Antischaum
Durchfluss
Desorber
Kolonne
Pumpe 324
Kopfdruck

Interface for context sensitive search terms

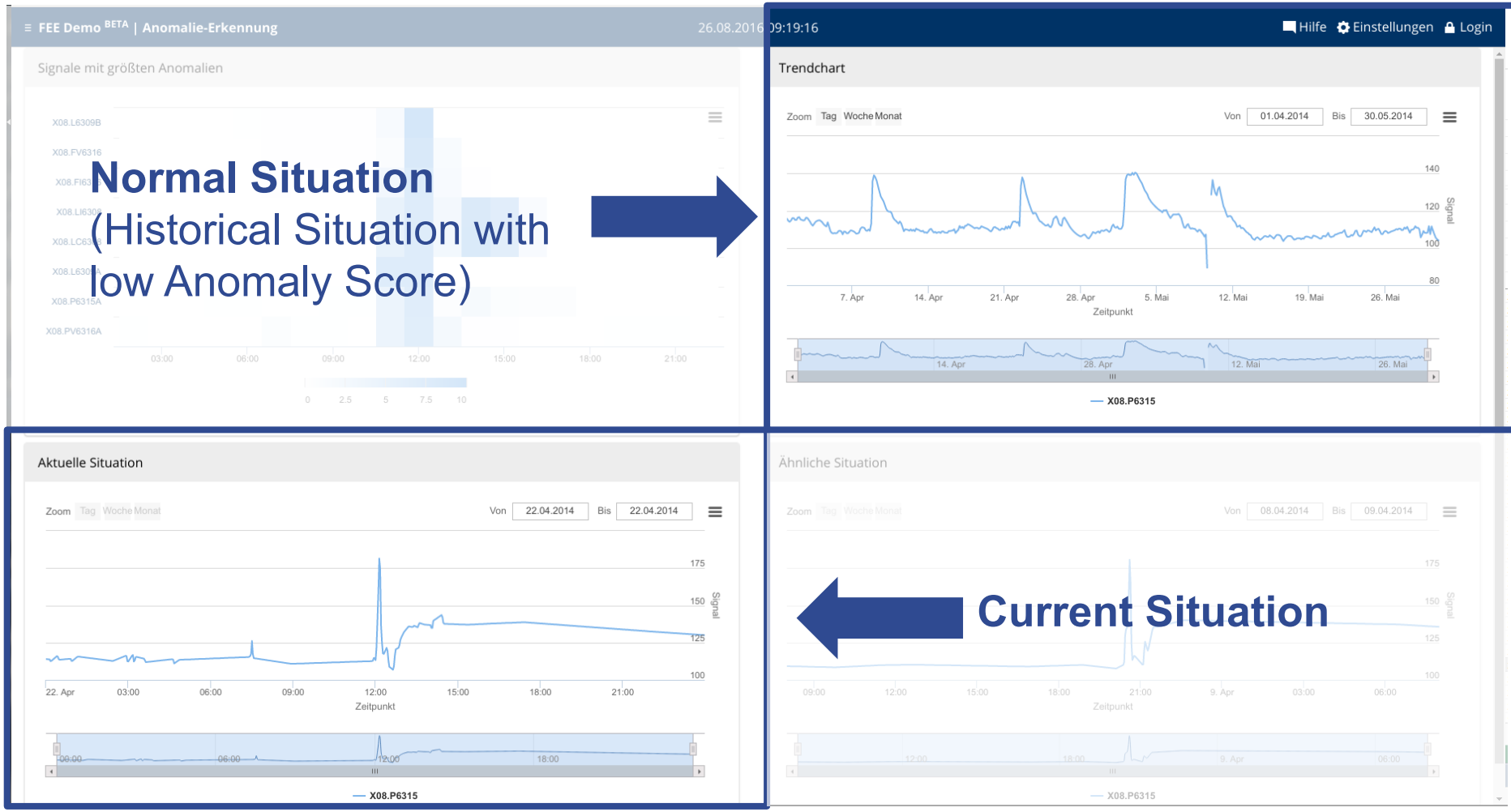


Operator Interface – Suspicous Signals (1)

Heat map:

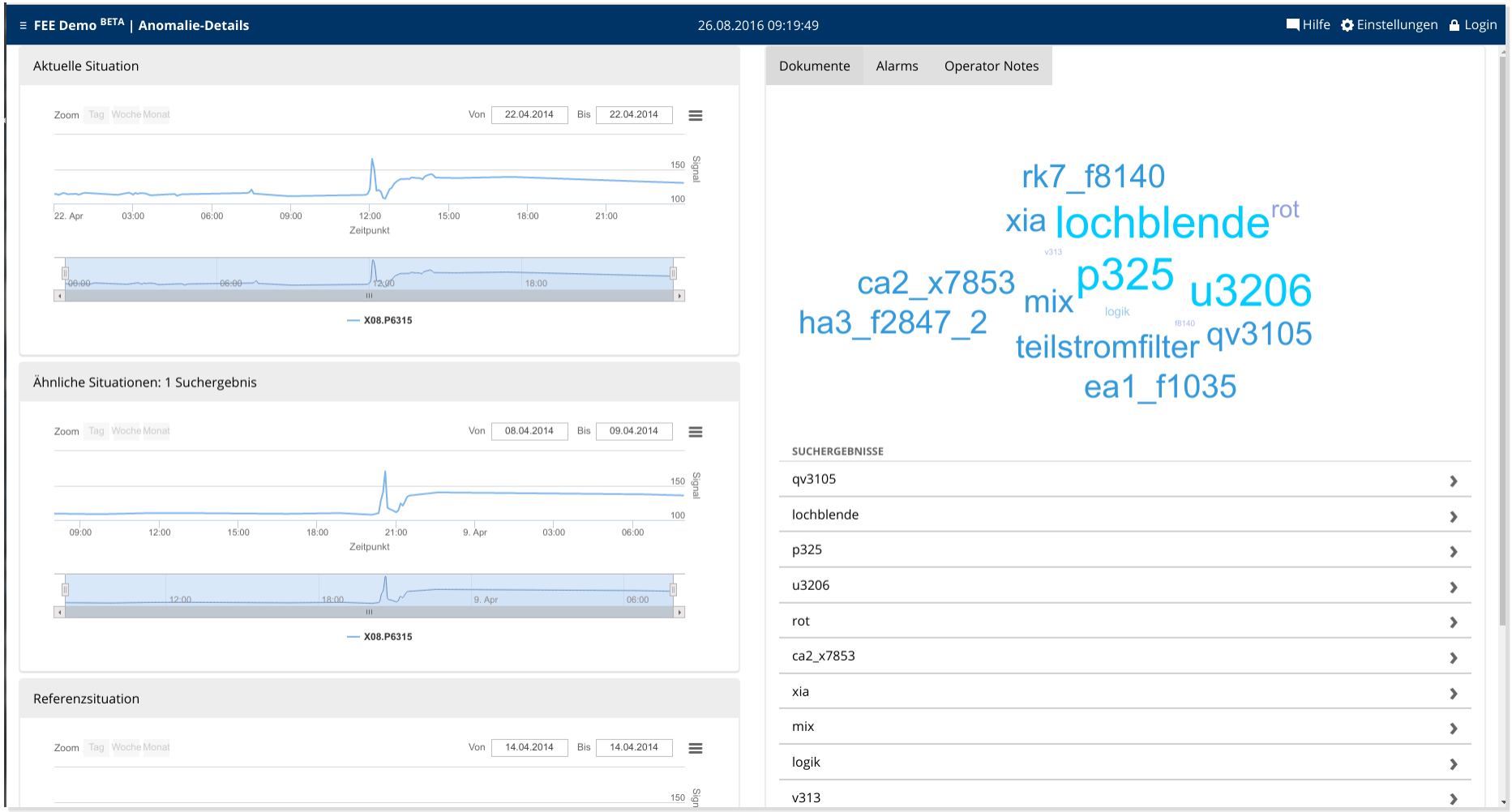
- Visualizes the signals with highest anomaly score
- Given an impression of the last couple hours
- Supports selecting single signals for detailed analysis

Operator Interface – Suspicous Signals (2)

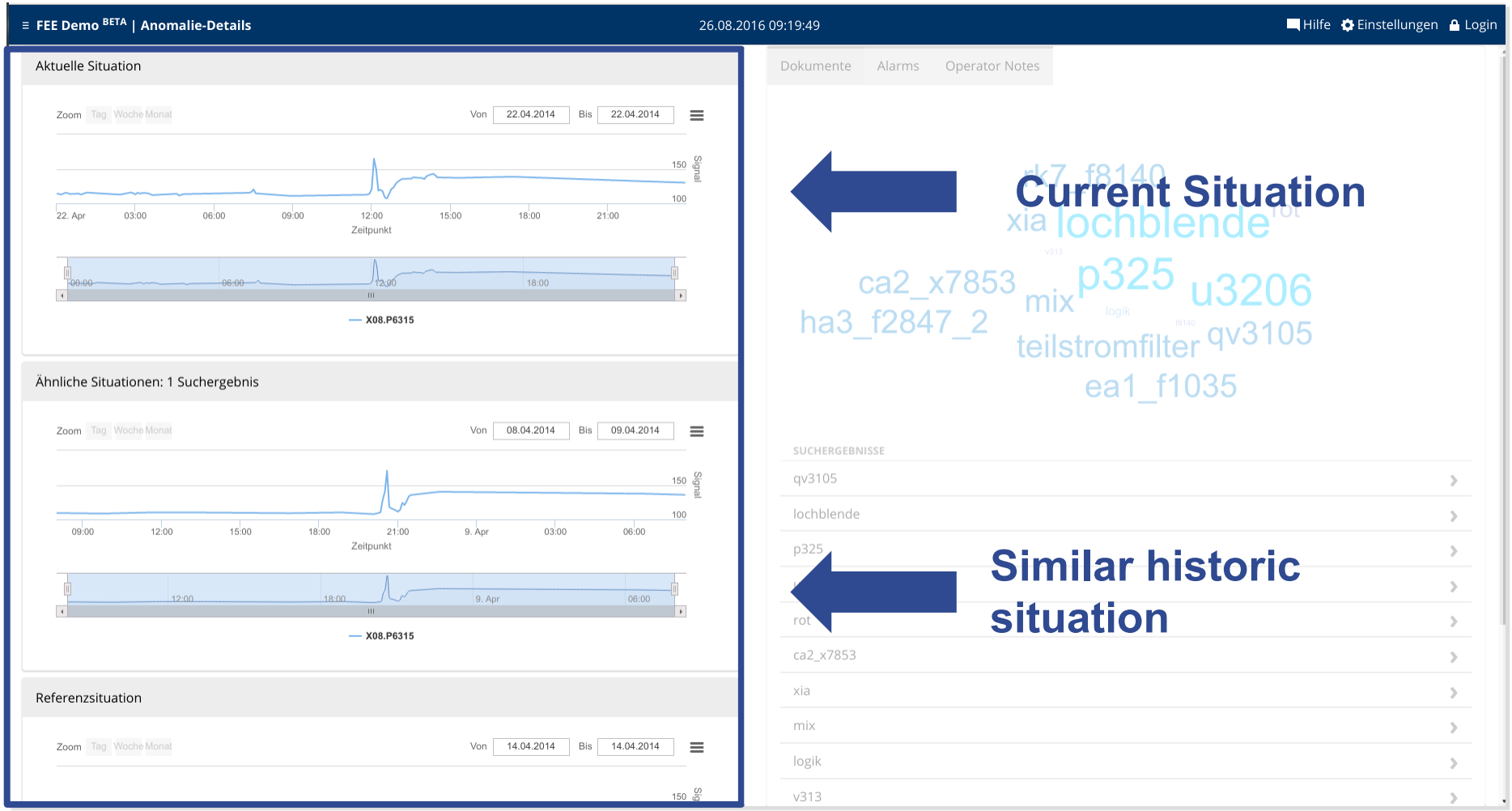


Operator Interface – Suspicous Signals (3)

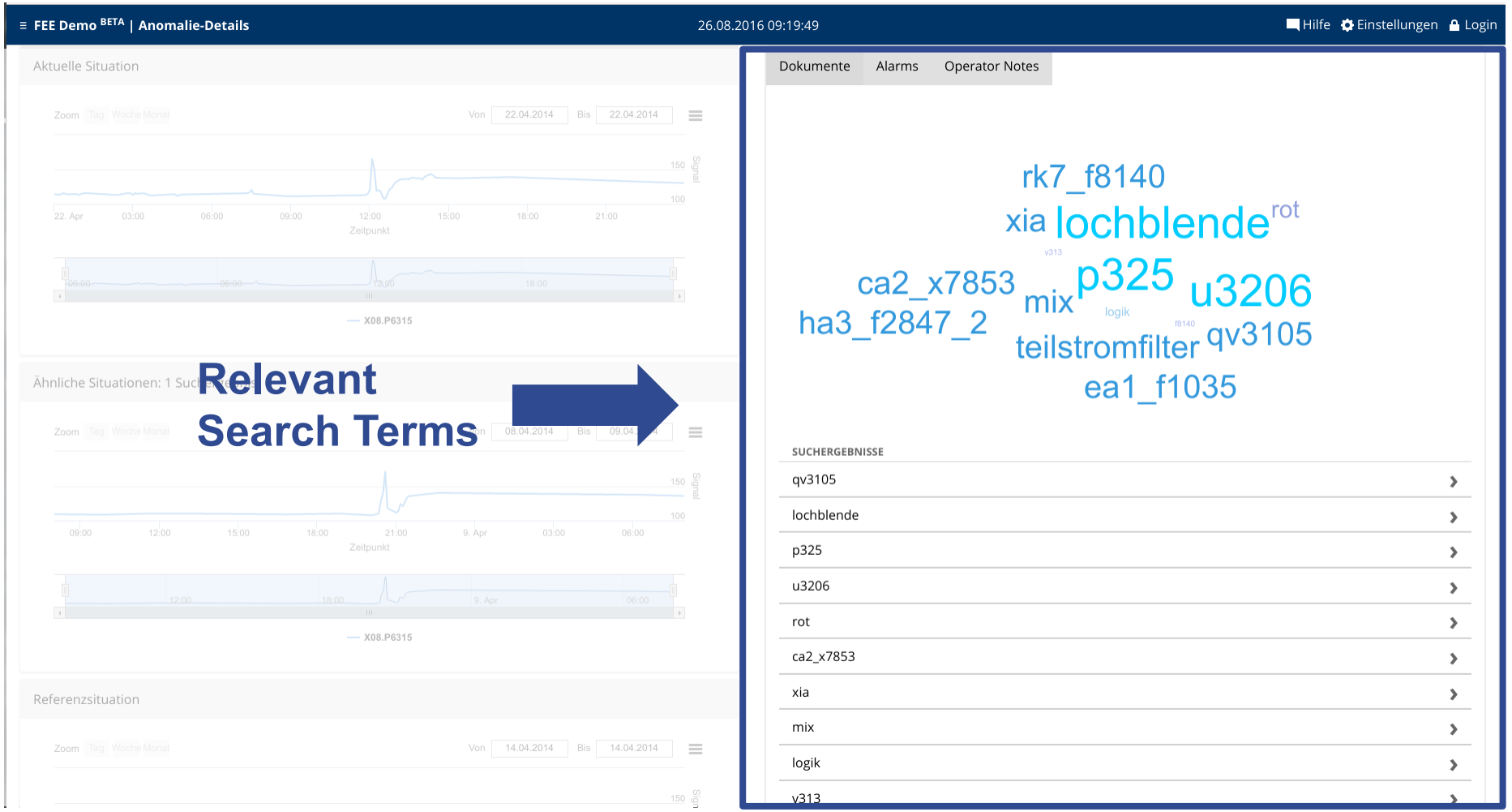
Operator Interface – Suspicous Signals (4)



Operator Interface – Suspicous Signals (5)



Operator Schnittstelle zur Anomalie-Erkennung (2)



FEE

Big Data for Operator Support in Chemical Plants

Summary

Deutsches Zentrum
DLR für Luft- und Raumfahrt e.V.
Projektträger im DLR

GEFÖRDERT VOM
Bundesministerium
für Bildung
und Forschung

Summary and Outlook

- What has been shown
  - Transfer of (big) data analytics into the context of chemical industry
  - Challenges of a big data architecture for chemical plants
  - Solution approach with two typical scenarios (Event prediction and anomaly detection)
- Next steps
  - Work on additional application scenarios
  - Further refinement of methods and demonstrating transfer to other plants
  - Demonstration of functionality in the plant context